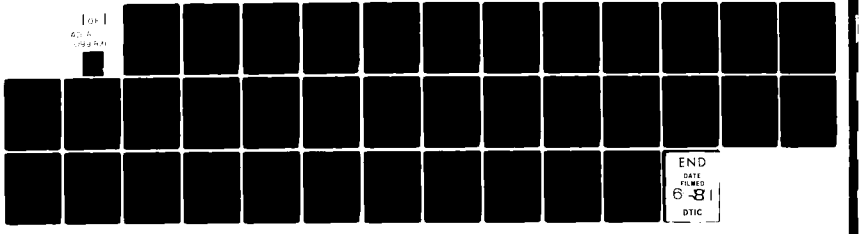
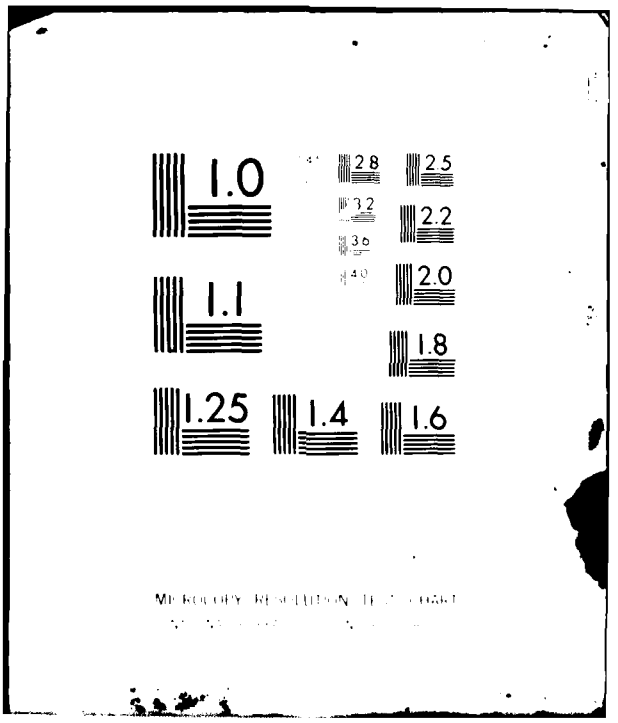
AD-A099 831 COURSEWARE INC SAN DIEGO CALIF F/G 5/9
QUALITY CONTROL PLANS, PROCEDURES, AND RATIONALE FOR THE F-16 P--ETC(U)
MAR 81 R R SUDWEEKS, A S GIBBONS, S J ROLNICK F02604-79-C-8875

UNCLASSIFIED NL




1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

LEVEL II

F-16 AIRCREW TRAINING DEVELOPMENT PROJECT

Contract No. F02604-79-C8875

QUALITY CONTROL PLANS, PROCEDURES, AND RATIONALE FOR THE F-16 PILOT TRAINING SYSTEM.

DEVELOPMENT REPORT No. 19
MARCH 1981

Prepared in fulfillment of CDRL no. B027,
and partial fulfillment of CDRL nos. B031 and B036

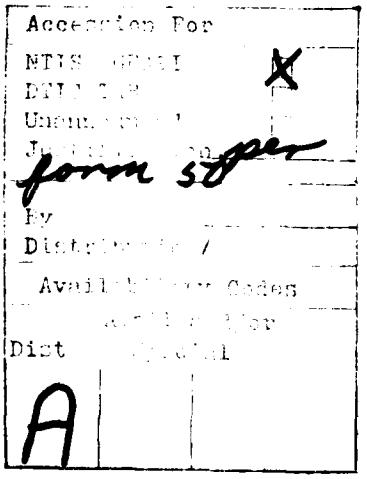
by

R.R Sudweeks
A.S. Gibbons
S.J. Rolnick
Courseware, Inc.

COURSEWARE, INC.
10075 Carroll Canyon Rd.
San Diego, CA 92131
(714) 578-1700

PREFACE

This report was created for the F-16 Aircrew Training Development Project contract no. F02604-79-C8875 for the Tactical Air Command to comply with the requirements of CDRL nos. B027, B031 and B036. The project entailed the design and development of an instructional system for the F-16 RTU and instructor pilots. During the course of the project, a series of development reports was issued describing processes and products. A list of those reports follows this page. The user is referred to Report No. 34, A Users Guide to the F-16 Training Development Reports, for an overview and explanation of the series, and Report No. 35, F-16 Final Report, for an overview of the Instructional System Development Project.



F-16 AIRCREW TRAINING DEVELOPMENT PROJECT REPORTS

Copies of these reports may be obtained by writing the Defense Technical Information Center, Cameron Station, Alexandria, Virginia 22314. All reports were reviewed and updated in March 81.

Gibbons, A.S., Rolnick, S.J., Mudrick, D. & Farrow, D.R. Program work plan (F-16 Development Report No. 1). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

Thompson, A., Bath, W., & Gibbons, A.S., Previous ISD program review (F-16 Development Report No. 2). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

Wild, M., & Farrow, D.R. Data collection and management forms report (F-16 Development Report No. 3). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

Gibbons, A.S. Review of existing F-16 task analysis (F-16 Development Report No. 4). San Diego, Calif.: Courseware, Inc., June 1977, March 1981.

Gibbons, A.S., & Rolnick, S.J. Derivation, formatting, and use of criterion-referenced objectives (CROs) and criterion-referenced tests (CRTs) (F-16 Development Report No. 5). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

Rolnick, S.J., Mudrick, D., Gibbons, A.S. & Clark, J. F-16 task analysis, criterion-referenced objective, and objectives hierarchy report (F-16 Development Report No. 6). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Gibbons, A.S. Task analysis methodology report (F-16 Development Report No. 7). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Gibbons, A.S. Objectives hierarchy analysis methodology report (F-16 Development Report No. 8). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Mudrick, D., Gibbons, A.S., & Schmidt, R.F. Goal analysis report (F-16 Development Report No. 9). San Diego, Calif.: Courseware, Inc., February 1978, March 1981.

Rolnick, S.J., Mudrick, D., & Thompson, E.A. Data base update procedures report (F-16 Development Report No. 10). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Mudrick, D., & Pyrz, K.E. Data automation of task and goal analysis: Existing system review and recommendation (F-16 Development Report No. 11). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

O'Neal, A.F., & Smith, L.H. Management System needs and design concept analysis (F-16 Development Report No. 12). San Diego, Calif.: Courseware, Inc., December 1977, March 1981.

Gibbons, A.S., Thompson, E.A., Schmidt, R.F., & Rolnick, S.J. F-16 pilot and instructor pilot target population study (F-16 Development Report No. 13). San Diego, Calif.: Courseware, Inc., September 1977, March 1981.

Schmidt, R.F., Gibbons, A.S., Jacobs, R. & Faust, G.W. Recommendations for the F-16 performance measurement system (F-16 Development Report No. 14). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Thompson, E.A., & Gibbons, A.S. Program/system constraints analysis report (F-16 Development Report No. 15). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Gibbons, A.S., & Rolnick, S.J. A study of media production and reproduction options for the F-16 project (F-16 Development Report No. 16). San Diego, Calif.: Courseware, Inc., February 1978, March 1981.

O'Neal, A.F., & Kearsley, G.P. Computer managed instruction for the F-16 training program (F-16 Development Report No. 17). San Diego, Calif.: Courseware, Inc., July 1978, March 1981.

Wilcox, W.C., McNabb, W.J., & Farrow, D.R. F-16 implementation and management plan report (F-16 Development Report No. 18). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Sudweeks, R.R., Rolnick, S.J., & Gibbons, A.S. Quality control plans, procedures, and rationale for the F-16 pilot training system (F-16 Development Report No. 19). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Gibbons, A.S., Axtell, R.H., & Hughes, J.A. F-16 media selection and utilization plan report (F-16 Development Report No. 20). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Thompson, E.A., Kearsley, G.P., Gibbons, A.S., & King, K. F-16 instructional system cost study report (F-16 Development Report No. 21). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Jacobs, R.S., & Gibbons, A.S. Recommendations for F-16 operational flight trainer (OFT) design improvements (F-16 Development Report No. 22). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Gibbons, A.S. F-16 instructional sequencing plan report (F-16 Development Report No. 23). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

Farrow, D.R., & King, K. F-16 coursewares and syllabi delivery schedule (F-16 Development Report No. 24). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Rothstein, L.J., Hibian, J.E., & Mudrick, D. F-16 instructor/course manager training requirements report (F-16 Development Report No. 25). San Diego, Calif.: Courseware, Inc., October 1978, March 1981.

O'Neal, A.F., & O'Neal, H.L. F-16 pilot media selection (F-16 Development Report No. 26). San Diego, Calif.: Courseware, Inc., March 1979, March 1981.

Gibbons, A.S. F-16 instructional system design alternatives (F-16 Development Report No. 27). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Gibbons, A.S. F-16 instructional system basing concept (F-16 Development Report No. 28). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

O'Neal, H.L., & Rothstein, L.J. Task listings and criterion-referenced objectives for the instructor pilot F-16 training program (F-16 Development Report No. 29). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Bergman, D.W., & Farrow, D.R. F-16 training system media report (F-16 Development Report No. 30). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Gibbons, A.S., O'Neal, A.F., Farrow, D.R., Axtell, R.H., & Hughes, J.A. F-16 training media mix (F-16 Development Report No. 31). San Diego, Calif.: Courseware, Inc. October, 1979, March 1981.

Farrow, D.R. F-16 training media support requirements (F-16 Development Report No. 32). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Gibbons, A.S. F-16 training media constraints and limitations (F-16 Development Report No. 33). San Diego, Calif.: Courseware, Inc., September 1979, March 1981.

Farrow, D.R., & Kearsley, G.P. A user's guide to the F-16 training development reports (F-16 Development Report No. 34). San Diego, Calif.: Courseware, Inc., January 1981, March 1981.

Farrow, D.R., & Clark, J. F-16 Final Report (F-16 Development Report No. 35). San Diego, Calif.: Courseware, Inc., January 1981, March 1981.

EXECUTIVE SUMMARY

Quality control is a fundamental component of the ISD process. In order to develop and maintain an efficient and effective training system, it is essential to monitor the system and make improvements on the basis of feedback data obtained from evaluation. Quality control procedures are applicable at three main stages in the development and implementation of a training system:

1. The developmental evaluation state (preimplementation).

2. The operational monitoring state (implementation).

3. Graduate evaluation (follow-up).

Common shortcomings of previous evaluation efforts in military ISD are:

1. Too narrow in scope.

2. Inappropriate data collected.

3. Incorrect analysis or interpretation of data.

4. Results are not utilized to rectify weaknesses in system.

The quality control plan for the F-16 project attempts to avoid these shortcomings. The developmental procedure will involve small-scale tryouts for the individual instructional segments followed by larger group tryouts with multi-segment materials. The operational monitoring stage will involve evaluation of data produced by the performance measurement system (PMS) which includes routine and systematic data collection procedures. The graduate evaluation stage will involve three types of data to be collected for each graduating class:

1. Task specific proficiency ratings.

2. Results of STAN/EVAL check ride, simulator check, and final examination.

3. Graduate evaluation questionnaires and structured interviews with graduates.

To assist in these three stages of quality control, evaluation specification checks are provided at six major checkpoints in the training cycle.

CONTENTS

LIST OF FIGURES AND TABLES

QUALITY CONTROL PLANS, PROCEDURES, AND RATIONALE FOR THE F-16 PILOT TRAINING SYSTEM

1.0 INTRODUCTION

This report contains a description of the evaluation activities to be employed in developing, installing, and maintaining the F-16 pilot training system. The training system includes an RTU course (in four different versions) and a continuation training course. The four versions of the RTU course are:

1. The F-16B (basic) course for pilots who have just completed Lead-in Training;

2. The F-16C (conversion) course for individuals converting from assignments as instructors in undergraduate pilot training (UPT) to assignments as F-16 pilots;

3. The F-16TX (transition) course for pilots who have already acquired mission-ready competencies in other aircraft; and

4. The F-16IP (instructor pilot) course for use in training pilots to be F-16 instructors.

Graduates of the B and C courses will be assigned to operational units where they will receive additional training and experience with the F-16 aircraft and weapons system by participating in the continuation training course.

The proposed plans assume that evaluation is an essential task in the development of effective training systems. The authors also contend that evaluation has a critical role to play in the implementation and utilization of a training system once it has been developed. The need to monitor operational systems and to make improvements on the basis of feedback data obtained from evaluating the functioning system and the performance of its graduates is emphasized. Without quality control procedures, the authors assert that instructional systems development (ISD) does not exist. These and other underlying ideas and assumptions are discussed in Section 2. Section 3 is a description of the various evaluation activities proposed in the F-16 project. Section 4 presents the schedule for conducting these activities

and Section 5 lists the personnel requirement and the responsibilities of the individuals involved. Evaluation specification sheets listing the issues and questions addressed at each of the six suggested quality checkpoints are found in the Appendix.

2.0 RATIONALE

ISD is a creative process which involves much more than making colorful overhead transparencies or multi-screen tape-slide presentations. ISD is a tool for analyzing training needs and then creating a complete system of instruction which meets those needs. The system may include (1) learning activities for the students, (2) teaching and evaluation procedures for the use of instructors, (3) supporting instructional materials such as workbooks, videotapes, or tape-slide presentations, and (4) implementation, management, and maintenance procedures for the system administrators. The components are intended to function together in a coordinated and complementary manner to consistently and predictably cause the students to acquire the skills, habits, understandings, attitudes, and judgment specified as intended outcomes. Instructional components which fail to produce the desired learner outcomes are not very useful in training situations even though they include fancy bells and whistles. An instructional system might well include entertaining activities and impressive materials, but if it is designed for specific training purposes, it should be judged in terms of its effects upon the intended learners rather than by its accoutrements.

2.1 Quality Control During the Development of Instructional Systems

Training and instruction are goal-directed endeavors in the sense that they are aimed at producing specified outcomes in a group of learners. Instructional events and conditions which can be repeatedly used to successfully produce the desired student outcomes are not accidents or fortuitous occurrences. They result from careful design and redesign. The generative process by which such events and conditions are devised involves forethought and planning, empirical testing, and then restructuring in the light of the test results. As part of the initial design process (1) the intended learner outcomes are pre-determined, (2) learning activities and materials are carefully planned and prearranged, and (3) procedures for the use of these activities and materials are prescribed in detail. These steps are done in the hope that the resulting instruction will yield the desired effects. However, until instructional events (procedures/activities/materials) have been tested and evaluated by observing their effects on learners there is no basis--other than theoretical predictions--for knowing that they yield the desired results. Theory and logic can be very helpful in the design process, but they must always be subject to the test of what happens in the real world. Accordingly, the effectiveness of a proposed instance of instruction must be verified through student-tryouts. To the extent that students acquire the desired learning outcomes as a result of using a particular instructional event, that event can be considered effective. The notion that instruction should

be evaluated in terms of its effects on the intended learner population is a fundamental idea in ISD.

To the extent that the instructional events are replicable and the results are measurable the instruction can be redesigned and progressively refined through a series of empirical tryouts and revisions. After each tryout of the instruction, an attempt is made to assess its effects on the learners. If the desired effects have not been obtained, adjustments are made in the procedures/activities/materials in an attempt to yield more desirable results. But the process does not end at this point. The revised instruction is then submitted to tryout and the results assessed. The tryout and revision cycle is repeated until the defects have been identified and eliminated and the desired results are obtained. This iterative process of design and redesign based on empirical tryouts constitutes a method of quality control.

The use of quality control procedures as an integral part of the process by which instructional systems are developed is an essential characteristic of ISD. Without quality control mechanisms, instructional systems development does not exist. Attempts to utilize ISD without quality control procedures are attempts to use a form without its substance. ISD is based upon the premise that instructional products (prescribed procedures, activities, or materials) are improvable. Each product is viewed as an approximation which can be refined and improved through the successive application of quality control procedures. Instructional products which have not been submitted to quality control procedures may have superficial characteristics of effective products, but they are not likely to consistently produce the desired results.

2.2 Quality Control in the Management of Instructional Systems

Evaluation is not conducted only during the development of instructional systems. The idea of improving system performance through the use of quality control procedures is just as applicable to the management of instructional systems as it is to their development. A training system produced by the ISD process is likely to include an assemblage of hardware and software components as well as detailed procedures for their use by trained personnel in an integrated manner. The hardware may include audiovisual-playback equipment as well as simulators and other sophisticated modeling devices. The software may include syllabi and other printed materials, videotapes and tape/slide presentations, plus forms and procedures for recording and evaluating pupil performance and progress. Since the various components are interdependent, to the extent that any of them are incompletely implemented or improperly utilized, the system is not likely to yield optimal results. The failure may be due to instructors who fail to follow prescribed procedures or it may be due to scheduling difficulties, inadequate equipment, or problems resulting

from the way the program is administered. Without a systematic procedure for monitoring the implementation and utilization of a training system, problems of this type are likely to remain undetected and uncorrected. It almost goes without saying, that it is foolish to spend large sums of money to develop a precision tool and then reap only part of the potential benefits just because the user was unaware that the tool was being improperly or incompletely utilized. Hence, there is an important need to carefully monitor the usage of a training system and its component parts.

Furthermore, the best laid plans of instructional developers do not always work well in the real-world contexts of Air Force training classrooms. The constraints under which a training system must be utilized at a particular location may reduce its effectiveness in one way or another. Careful monitoring of a system's operation through the use of quality control procedures should lead to the identification of any problems or deficiencies which exist. Necessary adaptations or modifications can then be consciously implemented to solve the problem. However, such problems are not likely to be corrected as long as they remain undetected.

2.3 Quality Control Tasks and Stages

Quality control in both instructional development and instructional management is essentially a troubleshooting or debugging process consisting of two main tasks: (1) diagnosing problems, and (2) prescribing solutions. The diagnostic task involves locating and correctly analyzing errors, omissions, malfunctioning strategies, and other weaknesses in the instructional components. The prescriptive task involves formulating revisions intended to solve the problems. Neither task is simple. Since unnecessary revisions are costly and counter-productive and since half-baked solutions may simply aggravate a problem, proposed revisions should not be made without good reason. Accurate diagnosis of instructional inadequacies requires skill in collecting and interpreting evidence. Apparent problems need to be verified and then carefully analyzed before any attempt is made to prescribe a solution. A clear understanding of the nature and source of the problem should lead to a higher probability of devising an adequate solution. Without a clear understanding of the problem, the proposed solution is not likely to solve the problem and the developers are likely to find themselves in the position of putting patches on ineffective patches.

Quality control procedures are applicable at three main stages in the development and use of a training system; (1) the developmental evaluation stage (2) the operational monitoring stage, and (3) graduate evaluation. As the chart in Figure 1 shows, the two main tasks--diagnosis and prescription--occur in each of these three stages. The entries in the cells of the

FIGURE 1. QUALITY CONTROL PROCEDURES

	Diagnose Problems	Prescribe Solutions
Developmental Evaluation	Review and tryout of materials in search of: 1. errors or omissions in the subject-matter content 2. weakness in the instruction 3. problems of style, grammer, or usage 4. inconsistencies between segments 5. errors in sequencing of tasks	Prescribe deletions, insertions, corrections, and other necessary revisions.
Operational Evaluation	Monitor utilization to assess: 1. degree of implementation and usage 2. problems and defects encountered by users 3. negative attitudes and other side-effects 4. inefficient or malfunctioning components	Prescribe alternative implementation and administrative procedures plus any necessary external or internal revisions.
Graduate Evaluation	Diagnose deficiencies in the performance of graduate pilots which can be attributed to inadequacies in their training	Prescribe corrections for each deficiency.

chart describe the tasks involved at each stage. The three stages are intended to be complementary, rather than redundant. Different questions and issues are addressed in each stage as a means of revealing different types of problems.

The developmental stage involves the development and refinement of prototype materials through an iterative process of reviews, tryouts, and revisions. This process is cyclical, but the number of iterations is typically small and the process is relatively short-lived. In contrast, the operational and graduate evaluation stages are used recurringly throughout the duration of the training system's use. Operational evaluation involves continuous monitoring of the implementation and utilization of the system in an attempt to identify and correct defects in the management procedures or problems in the context in which the system is used. Graduate evaluation is a follow-up procedure which occurs in the succeeding months after each class completes the course. This third phase involves appraisal of graduates' on-the-job performance in operational units. The intent is to discover any deficiencies in the pilots' performance resulting from inadequacies in the training system. More specifically, it is an attempt to identify areas in which the training experienced by the pilots may have been insufficient, irrelevant, or otherwise inadequate.

Each stage of the quality control process is a formative evaluation procedure. Each stage includes one or more corrective feedback loops as graphically depicted by the flow chart in Figure 2. Prescribed modifications, replacements, deletions, or other revisions are fed back into the system at the appropriate point as a means of improving it. Changes in management practices prescribed at the second or third stages are inserted in at the operational stage, but revisions prescribed as a result of deficiencies in the instruction must be submitted to developmental tryout. Together, the three stages provide a systematic means of progressively refining both the training system and the management practices by which it is implemented and utilized. The overall purpose is to maximize the effectiveness of the system. The overriding concern is to accurately diagnose and correct deficiencies in the system at a minimum cost in terms of time, personnel, dollars, and interference with on-going Air Force operations.

2.4 Common Shortcomings of Evaluation Efforts in Military ISD

Attempts by instructional developers to apply quality control procedures in military contexts are frequently characterized by one or more of the following shortcomings:

1. The scope of the evaluative effort was too narrow. Important questions and issues were overlooked or ignored.

2. Inappropriate or insufficient data were collected.

3. The data collected were not analyzed completely or were incorrectly interpreted.

4. The results were not utilized to correct weaknesses and deficiencies in the instructional system.

The first problem occurs if one or more stages of the quality control process is omitted. For example, if no attempt is made to monitor the implementation and utilization of a system, important contextual factors impinging upon its effectiveness will likely be overlooked.

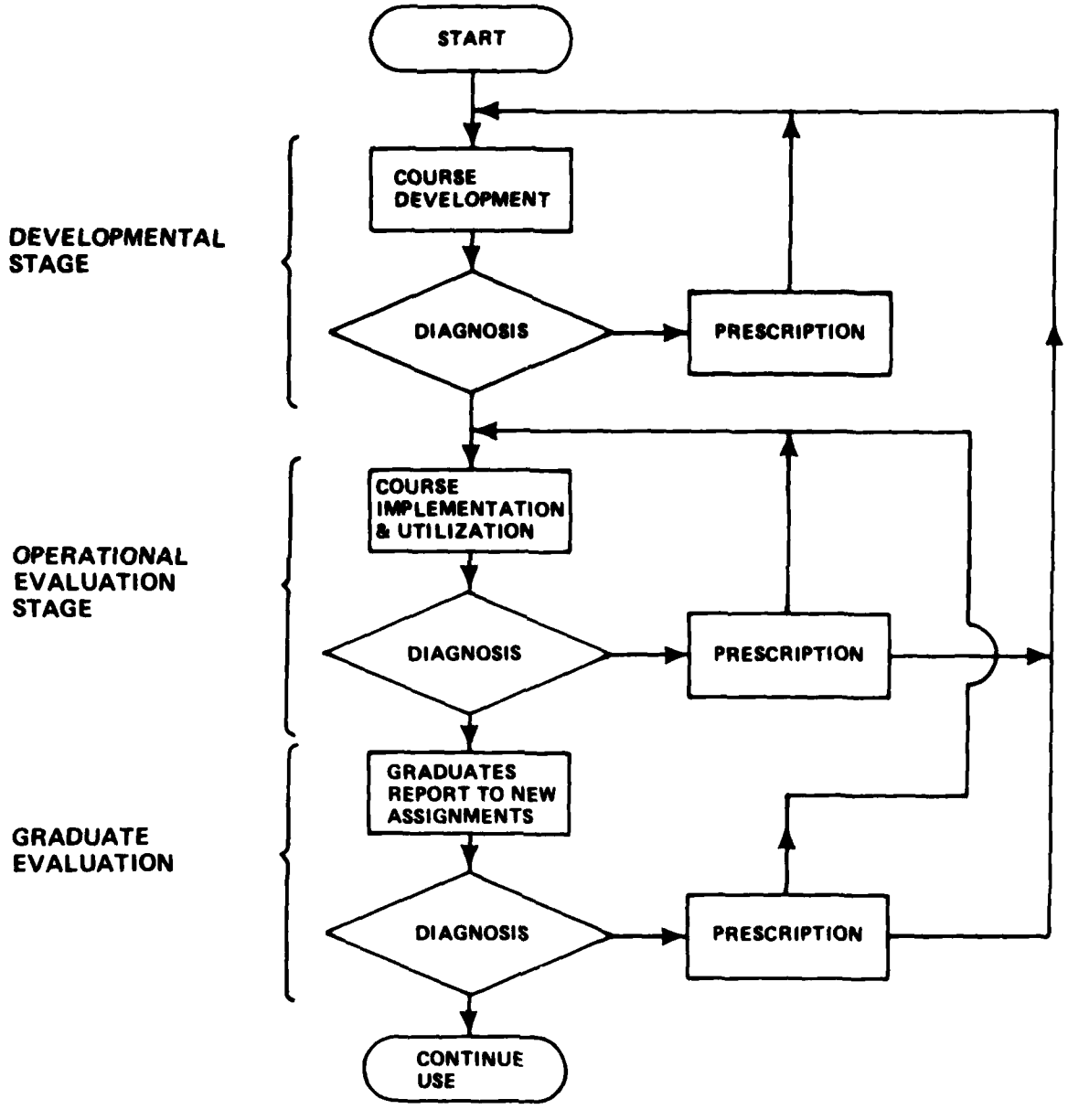
The second problem occurs when the evaluation design is inadequately planned or poorly executed. This problem can result from failure to keep adequate records or from the use of invalid or unreliable instruments to collect data. The low rate of return of questionnaires used in the Air Force graduate evaluation program is an example of insufficient data being collected. In absolute terms, the amount of data collected would probably be sufficient if there were legitimate grounds for assuming that the proportion returned was representative of the whole population. However, representativeness is not likely in this situation and can not justifiably be assumed without the use of probability sampling procedures.

The use of impeccable instruments and procedures for collecting the appropriate data do not insure that the data will be completely analyzed and correctly interpreted. In fact, piles of data collected in the operational and follow-up phases of military projects often go unanalyzed because of changes in priorities and/or lack of resources for carrying out the analysis. Unanalyzed data are not helpful and represent a waste of time and money. Erroneous interpretations or conclusions drawn from data are more likely to be misleading than helpful in diagnosing and correcting instructional problems.

The fourth problem occurs when the operations training development (OTD) team or other personnel responsible for revising a system do not have access to the results of an evaluation or do not use them for one reason or another. The results of proficiency checks and written exams administered by STAN/EVAL teams is a contemporary example of this problem. The Air Force could get more mileage out of the dollars spent administering these performance checks if the data were made available to OTD teams.

The practical result of all four problems is that the feedback loop is either incomplete or non-functional. The whole notion of progressive refinement in the development and management of instructional systems is based on the use of feedback information and judgments in the revision-making process. To the extent that the feedback loops are non-functional, quality control does not exist.

FIGURE 2. THE THREE STAGE QUALITY CONTROL PROCESS



3.0 QUALITY CONTROL PLANS FOR THE F-16 PROJECT

The plans for evaluating the F-16 Project are an attempt to utilize the ideas presented in Section 2. A deliberate attempt has been made to avoid the shortcomings described in Section 2.4. F-16 evaluation activities are designed to occur in the three stages delineated in Section 2 and displayed in Figure 1. Detailed explanations of the activities planned for each of the three stages are presented in the remainder of this section.

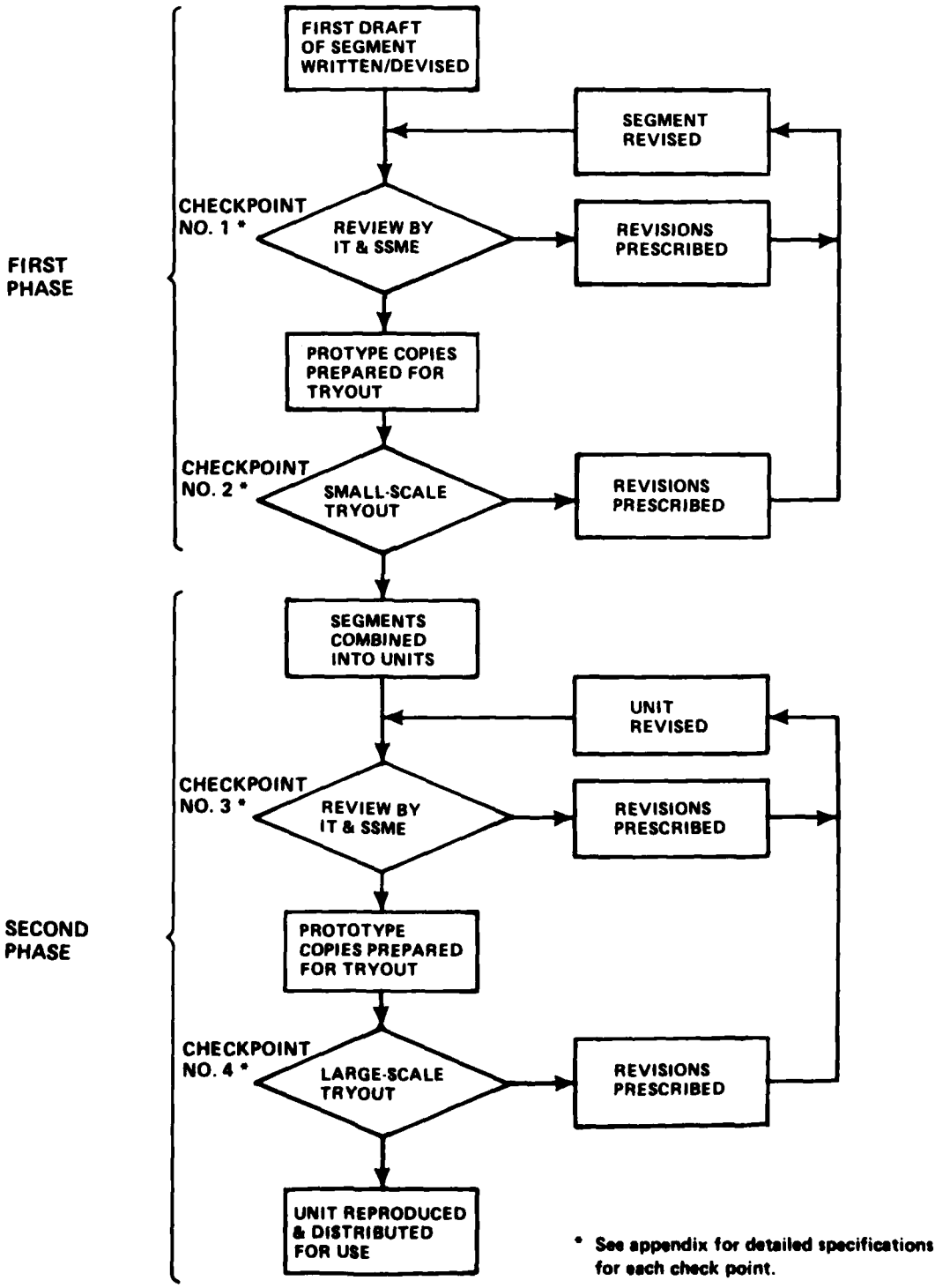
3.1 The Developmental Stage

Interesting and effective instructional components do not spring from the mind of an author in a complete and perfect state. Instead, a rough approximation of the desired component is first produced and then refined through a programmatic, empirical process of identifying and eliminating defects. The procedure is programmatic in the sense that (1) there is a commitment to producing the predetermined learner outcomes, and (2) strategies which are demonstrated to be ineffective as a means of accomplishing the desired outcomes are abandoned or replaced with other approaches which hold promise of being more effective. The procedure is empirical in the sense that decisions about the effectiveness of a strategy are determined by its observed effects on students.

The developmental procedure that is to be used in the F-16 project includes two main phases. The first phase involves small-scale tryouts of individual segments of instruction. The segments are combined into units in the second phase and tried out by larger groups of students. Each phase includes both a review and a tryout step. The sequence of steps in each phase is shown in the flow chart in Figure 3. The four diamonds in the flow chart represent quality control checkpoints. The issues and questions to be addressed at each checkpoint are listed on the corresponding evaluation specification sheets found in the Appendix.

Draft copies of instructional segments or training session plans will be initially written by subject-matter experts (SMEs) from the Air Force F-16 OTD team using design specifications prepared by instructional technologists and instructional psychologists employed by the contractor. Each draft copy will be reviewed by two different individuals, first by an instructional technologist and then by a senior subject-matter expert (SSME). The instructional technologist will check to see that the segment is complete and that it complies with the design specifications. The instructional technologist will also be responsible for editing the segment in terms of style, grammar, usage, and format. The Air Force SSME will inspect the technical accuracy of the subject-matter content. On the basis of their findings, the instructional technologist and SSME will be responsible for

FIGURE 3. THE DEVELOPMENTAL EVALUATION PROCESS



* See appendix for detailed specifications for each check point.

preparing revision specifications which indicate the number, type, and location of all changes needed. Simple revisions may be made on the spot by either reviewer. After the segment has been revised by the authoring SME, it will be briefly checked again by the SSME and instructional technologist to verify that the revisions have been made in accordance with the specifications indicated.

Prototype copies of the amended segment will then be prepared for tryout. All instructional components of the segment are to be included. However, visuals and audio materials should be presented in a rough, unfinished form which adequately conveys the message but can be produced relatively inexpensively. A small sample of three to seven students typical of the target population will be selected to try out the segment. The students will be asked to study the materials and to make note of ambiguous and/or confusing portions of the material as well as overly-simple or overly-difficult sections. Usage of the materials by the student-pilots will be carefully observed in an attempt to discover any problems they encounter. An appropriate criterion measure will be administered when the students complete the instruction. The results will be diagnosed to determine what the students have and have not learned and what revisions need to be made to improve the students' level of mastery. The materials may then be recycled and tried out again following incorporation of the specified revisions or they may be set aside until the beginning of the next phase in the developmental stage.

The second phase of development involves combining several instructional segments into one unit and then trying them on a larger group of students under typical training conditions. After completing the instruction the students will be tested and their performance carefully analyzed in search of misconceptions and other problems due to any inadequacies or inaccuracies in the materials. Student affect towards the materials and procedures will also be investigated at this time in an attempt to identify sources of negative attitudes and other unintended affective outcomes. Students will be asked to identify any confusing or ambiguous directions. If necessary, the materials should be revised further and then recycled through another tryout phase. Otherwise, they should be sent to a production instructional technologist in order to have press-ready or master copies prepared for production and distribution.

Since prototype components are easier and less expensive to correct than finished products, extra caution will be taken during the developmental stage to identify and correct as many content errors and instructional defects as possible. Astute thinking and painstaking attention to detail during this stage should save hours of time, stacks of paper, and hundreds of dollars in layout and production costs later. Once the instructional components have passed through this first screening stage, they should be free from all but the most-difficult-to-detect

mistakes and defects. They are then ready to be mass-produced and distributed for adoption and use.

3.2 The Operational Monitoring Stage

Operational evaluation addresses issues related to the way in which the training system is installed and used in a particular location. It involves continuous monitoring of system performance in an attempt to identify defects which impair the efficiency or the utility of the system. During developmental evaluation the major concern was to assess and improve the effectiveness of instructional strategies and prototype components. In the operational stage the focus shifts from internal to external characteristics of the system. There is still a concern for effectiveness, but primary attention is given to administrative factors and contextual contraints which affect the system and the manner in which the components are utilized.

Operational evaluation is a management and control tool. It consists of a set of evaluation and reporting procedures which should routinely be employed each time the course is taught at each training site. Since the context in which the course is utilized will vary from place to place, results of the operational evaluation will likely be site-specific. The purpose is not to produce conclusions which can be generalized across sites so much as it is to identify and correct problems at each installation in an attempt to maximize the efficiency and utility of the system. Regardless of whether a problem is common to all sites or peculiar to one, it needs to be rectified if it is inhibiting the system from functioning in an optimal manner.

Specific information regarding the issues to be investigated in this stage is presented in the evaluation specification sheet for Checkpoint #5 (see Appendix). Usage of the various materials, equipment, and training devices will be monitored at each site to determine to what extent the system is being utilized and the degree of adherance to implementation plans. The environment in which the system functions will be assessed to determine to what extent the prescribed conditions prevail at each site and what limitations are imposed by contextual constraints. The flow of students through the system will be monitored as a means of assessing adherance to schedule and as a means of identifying bottlenecks in the system which need attention.

Assessment of the extent to which the performance measurement system and the system update procedures are being utilized should receive special attention during operational evaluation. Since the F-16 is an emerging weapon system it is anticipated that the RTU course will need to be revised and updated as procedures for operating the F-16 evolve. The various system update procedures are designed for use in keeping the task listings, criterion-referenced objectives, and subject-matter content current and up-to-date as changes in F-16 procedures are initi-

ated by the Air Force. The degree to which these maintenance procedures are being used and any difficulties encountered should be reported as part of the operational evaluation.

The performance measurement system (see Project Report No. 14, "Recommendations for the F-16 Performance Measurement System") provides a systematic means of routinely collecting and assessing student performance data from the various aspects of the training system. The data generated by this system provides evidence of the relative effectiveness and or ineffectiveness of various components in the training system. Usage data must be consulted for any products which are shown to be lacking in effectiveness. It may be that such components are not being utilized in the intended manner.

Interview data from students, instructors, and project administrators will be collected. Instructors will be interviewed to determine how satisfied they are with the system, what they perceive its strengths and weaknesses to be, and what common or recurring problems they confront.

Students will be given an opportunity to express their feelings about the utility, relevance, and difficulty of the training system. They will also be asked to identify sections of the course where they need more (or less) time. Some training experiences may not be maximally effective simply because sufficient time was not provided.

3.3 Graduate Evaluation

Graduate evaluation is the process of reviewing assessments which have been made of the graduates' on-the-job performance in search of common problems and frequently-occurring deficiencies. It is a retrospective activity which takes place on a recurring basis in the succeeding months after each group of students complete the training course. It is an attempt to identify deficiencies in the performance of the graduate pilots which can be attributed to inadequacies in their training.

The purpose of graduate evaluation is not to grade or reward the pilots, but to improve the training system. It is the training system that is on trial, not the pilots. The two main issues are (1) What necessary competencies, proficiencies, or attitudes have the pilots not yet acquired?, and (2) What unnecessary or irrelevant skills or attitudes have the pilots acquired which they would be better off without?

Since careless errors, bad habits, and poor judgment can all have dire, if not fatal, consequences for an F-16 pilot; it is not enough just to determine what competencies a pilot has or does not have. The manner in which he customarily uses his competencies must also be examined in search of high-risk mistakes, bad habits, and evidence of poor judgment. Consequently,

there must be some provision for unobtrusively assessing the pilot's usual performance in the routine tasks of operating the aircraft and weapon system when he does not know he is being rated or evaluated on a particular set of tasks or procedures.

The specific issues and questions to be addressed in the follow-up evaluation are listed in Checkpoint #6 in the Appendix. Three types of data will be collected and reviewed for each graduating class:

1. Task-specific proficiency ratings recorded on gradeslips by IP's during continuation training;

2. Results of the first STAN/EVAL check ride, simulator check, and written exam administered to each pilot after RTU graduation;

3. Responses to open-ended questions collected from structured-interviews by telephone with a systematic sample of graduates.

One purpose of continuation training is to provide RTU graduates with additional experience both in the aircraft and in a simulator. During continuation training each student's performance on specific tasks will be rated by an instructor pilot using a gradeslip. Copies of the first 12 gradeslips completed for a student when he begins continuation training will be to be returned to the F-16 OTD team for review and analysis. The student's flight commander will be held responsible for ensuring that the gradeslips are sent.

TACR 60-2 specifies that flight and instrument proficiency checks plus a written exam be administered to all operational pilots twice yearly by an independent Air Force evaluation group organized for the purpose. The flight and instrument checks are based on checklists completed by an observer-evaluator during and after the check-ride. The written exam covers all procedures. The resulting ratings, performance deviations, and comments are summarized on TAC Form 8.

The STAN/EVAL assessment data is already being generated by existing procedures, but at present the results are used for TAC evaluative purposes only and are not made available to OTD teams for diagnostic purposes. It is the recommendation of the present evaluators that STAN/EVAL data for graduates of the F-16 RTU courses be made available on a regular basis to the F-16 OTD team. The returns could be anonymous if necessary, since it is not the students, but the training course that is being evaluated by the OTD team. The criteria used by the STAN/EVAL teams should be based upon the criterion-referenced objectives in the F-16 course. Otherwise, the pilots will be evaluated in terms of two different sets of criteria.

Graduate pilots are valuable sources of information regarding the relevance, completeness, and adequacy of the course. If a pilot received irrelevant, incomplete, or otherwise inadequate training, by the time he has been in continuation training for three months it is likely that he will have identified some of the weaknesses in his RTU experience. He should know the areas in which he was well-trained as well as at least some of the tasks in which his training was weak or skimpy. About three months after each group of pilots have graduated from the F-16 RTU course, a probability sample of the graduates should be interviewed by telephone and queried about the relative adequacy of the F-16 RTU course in view of their subsequent experience.

All members of the first two graduating classes from each training site will receive follow-up interviews. A systematic sample of one-half of the graduates will be selected from subsequent classes. Once the training course has become well-established at a particular site the sampling fraction could be reduced to one-third, but whenever major changes are introduced into the course the fraction should be increased. The sample will be selected by listing the names of the graduates from a given class in the order of their class standing and then selecting every second (or third) name after randomly choosing the starting point in the first interval. The use of this procedure will insure that the sample includes pilots who did very well in the RTU course as well as pilots from the lower end of the class.

3.4 Maintenance Development

The operational monitoring and graduate evaluation procedures are intended to be used recurringly in connection with each use of an F-16 course, but the initial developmental stage is intended to be a one-time activity. However, because of the emerging nature of the F-16 weapon system and because some the "developed" components may possess some serious defects, it is anticipated that periodically during the life-cycle of the instructional system it will be necessary to create additional or replacement components or to further refine some of the existing components. These new components should be submitted to the same process of developmental tryouts and revisions as the original components were. This occasional--as needed--use of developmental procedures is a process of maintenance development (see Figure 4) and is a very important means of retarding the obsolescence of the instructional system. The system update procedures for updating the task listings, criterion-referenced objectives, and subject-matter content should always be used in conjunction with such new developmental activities. Other data from the operational and graduate evaluation stages may well be consulted also when using this maintenance procedure.

FIGURE 4. MAINTENANCE DEVELOPMENT

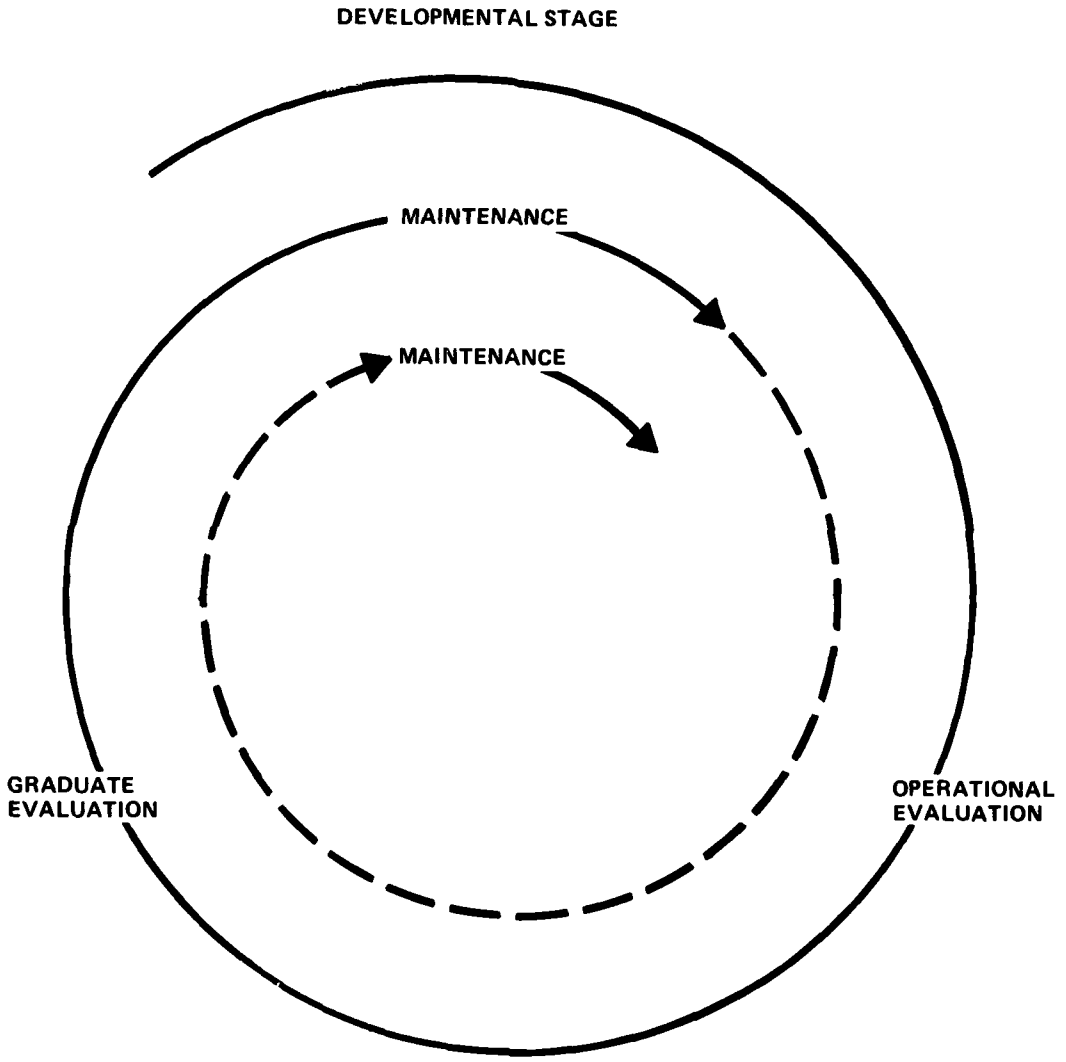
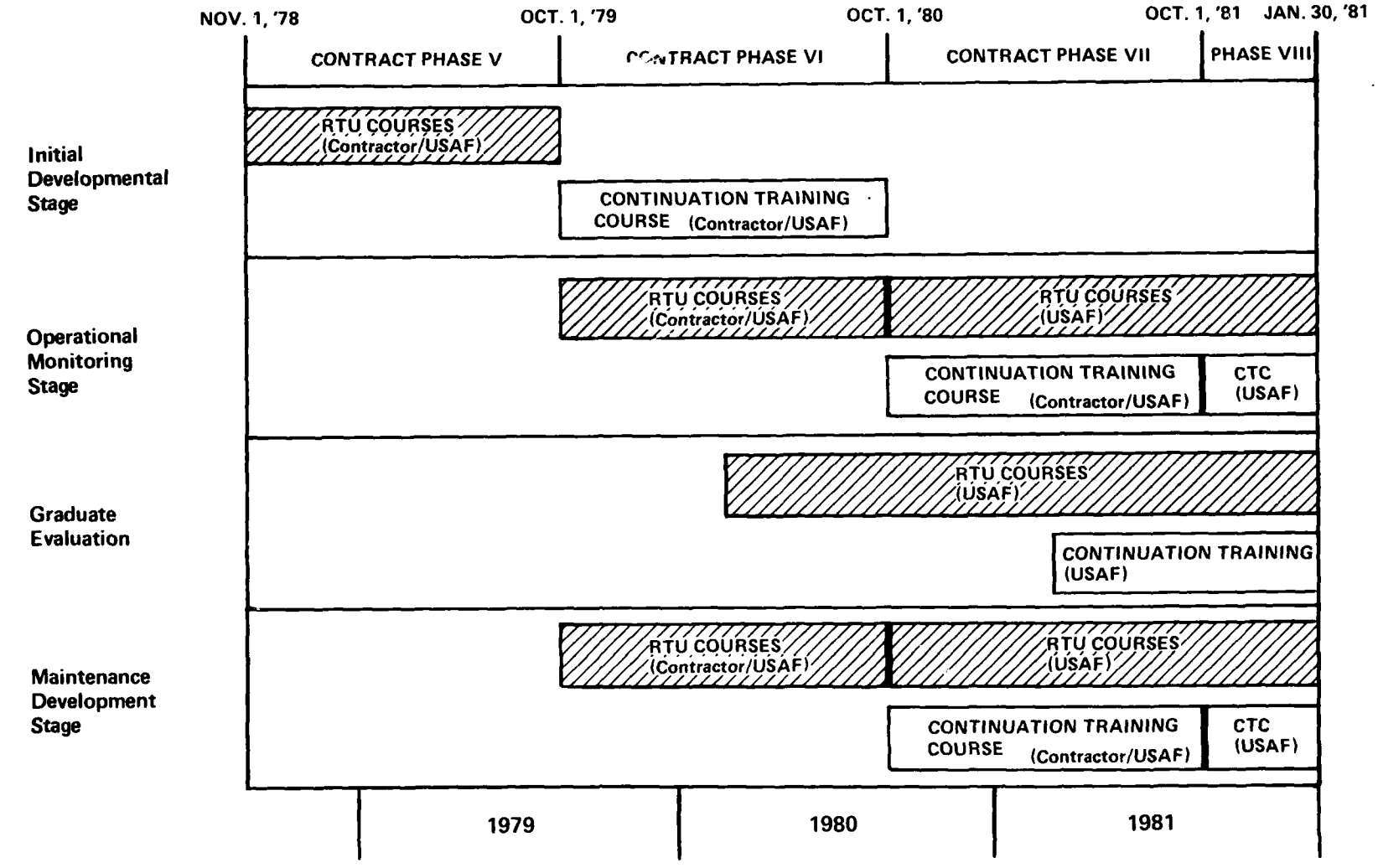


FIGURE 5. SCHEDULE OF QUALITY CONTROL ACTIVITIES



4.0 SCHEDULE

The chart in Figure 5 displays the schedule of quality control activities in the F-16 project plotted accross time. During the year between November 1, 1978, and October 31, 1979, the three RTU courses will be produced by means of the developmental process previously described in this report. Operational monitoring will begin when implementation and utilization of the course are started. Assuming that the duration of the course is approximately six months, graduate evaluation would begin soon thereafter. The maintenance procedure will be employed on an as needed basis whenever the results of either the operational monitoring or graduate evaluation procedures indicate the need to develop additional or substitute components to be included in the course. Initially, the operational monitoring and maintenance development procedures will be accomplished through a joint effort also. However, after the RTU courses have each been used twice, the contractor will phase out of these two evaluation activities and leave the OTD team to continue them. Note, that the contractor is not scheduled to assist in graduate evaluation. This will be the OTD team's responsibility right from the beginning. The same schedule will be followed in evaluating the continuation training course, except that it is delayed for one year until the initial development of the RTU course is completed.

TABLE 1. PERSONNEL REQUIREMENTS

Major Tasks	Person Responsible for Collecting Data	Person Responsible for Analyzing Data	Frequency
Developmental Stage			
Content Reviews	Senior SME	Senior SME	At least once for each instructional
Style/Format Reviews	IT	IT	
Small-scale Tryouts	IT/IP	IT/IP	
Large-scale Tryouts	IT/IP	IT/IP	
Operational Stage			
Describe and judge system performance with checklist	ISD Monitor in each wing	OTD Team Evaluation (contractor)	Twice per class per site
Interview instructors	ISD Monitor in each wing	OTD Team Evaluation (contractor)	Once per class per site
Graduate Evaluation			
Collect & Analyze gradeslips	Wing	OTD Team Evaluator (contractor)	Once per site per year
Collect & Analyze STAN/EVAL proficiency ratings of graduates	STAN/EVAL Teams	OTD Team Evaluator	Once per year
Conduct telephone interviews with selected graduates	OTD Team Evaluator	OTD Team Evaluator (contractor)	Once sample per class per site

5.0 PERSONNEL REQUIREMENTS AND RESPONSIBILITIES

Successful execution of this F-16 quality control plan for the F-16 pilot training system will require use of personnel supplied by both the Air Force and the contractor. The tasks to be performed in each stage of the evaluation are listed in the left column of Table 1. The persons responsible for performing each of these tasks are listed in the two middle columns. The Air Force will be expected to supply two SSMEs during the developmental stage. In the operational stage, the Air Force will be expected to appoint a responsible person at each training site to monitor the system's operation using the checklist provided and to interview the instructors. Since TACR 50-1 (paragraph 8) stipulates that each training wing or group appoint an ISD Monitor, the responsibility of monitoring system operation should probably be assigned to that individual. The monitor in each wing will be responsible for sending the data to the F-16 OTD team. One member of the OTD team should be assigned as the evaluator with the responsibility of analyzing the data collected from the various training sites and reporting the results to the OTD team commander.

Graduate evaluation is based on data collected from three sources: (1) gradeslips completed by instructor pilots for each pilot in continuation training, (2) proficiency checks by STAN/EVAL teams, and (3) telephone interviews with graduate pilots. Each continuation training wing should be responsible for collecting the first 12 gradeslips given each pilot and forwarding them to the OTD team. The evaluator on the OTD team has the responsiblity for analyzing, summarizing, and reporting the results of the gradeslips, proficiency checks, and interview data. He is also responsible for conducting the student interviews.

As a unit the OTD team has the responsiblity of maintaining and improving the system. It is their responsibility to interpret the evaluative results and to make specific recommendations for revising system components and implementation plans. Representatives from the OTD team should make site visits as necessary to the various training sites in order to further investigate problems needing their attention. All recommendations made by the OTD team should be forwarded to DOOS as well as to the commanders of the respective units.

Appendix

EVALUATION SPECIFICATION SHEETS

Checkpoint #1

Segment Review by Instructional Technologist
and Senior Subject-matter Expert

EVALUATION OBJECT:

Draft Copies of Instructional Components and Training Session Plans

ISSUES/QUESTIONS TO BE ADDRESSED:

1. Technical accuracy of the subject-matter
   a. What errors or omissions need to be corrected?
   b. Which sections need to be up-dated? In what ways?
2. Instructional soundness
   a. In what ways have the instructional specifications not been met?
   b. How adequate is each of the instructional components in terms of generalities, helps, example sets, and practice items?
   c. To what extent are the presentations consistent with the objectives and test items?
   d. Which subskills have not been included?
   e. Have common mistakes identified in the tryout of previous segments been avoided or adequately provided for? Which ones have not been?
3. Style
   a. What spelling, punctuation, grammatical, or typographical errors exist?
   b. In what ways are the vocabulary or idiom inappropriate?
   c. How can the clarity and tone of the text or narration be improved?

INSTRUMENTATION:

Checklist

Checkpoint #2

Small-scale Tryout of Prototype
Instructional Segments

EVALUATION OBJECT:

Prototype Versions of Instructional Segments

ISSUES/QUESTIONS:

1. Effectiveness of the segment
   a. To what extent do students attain mastery of the desired behaviors as a result of the instruction?
   b. Which portions or sections of the materials are ineffective or malfunctioning? In what way?
2. Student reactions to the material
   a. Which sections were perceived by the students as being too difficult or too easy? Why?
   b. Which sections were perceived as being uninteresting to the students? What suggestions do students have for improving the interest level?

INSTRUMENTATION:

Mastery tests
Interviews with students

Checkpoint #3

Unit Review by Instructional Technologist
and Senior Subject-matter Expect

OBJECT OR EVENT TO BE EVALUATED:

Instructional Segments Revised and Combined into Units

ISSUES/QUESTIONS TO BE ADDRESSED:

1. Technical accuacy of subject-matter
   a. What errors or omissions need to be corrected?
2. Instructional soundness
   a. Are any segments missing or incomplete?
   b. To what extent is the sequence of the segments consistent with the task listing and objectives hierarchy? How can the sequence be improved?
   c. In what ways are segments in the unit inconsistent or incompatible?
3. Style
   a. Are all of the materials (including charts and diagrams) legible and readable?
   b. What spelling, punctuation, grammatical, or typographical errors exist?

INSTRUMENTATION:

Checklist

Checkpoint #4

Large-group Tryouts of Instructional Units

EVALUATION OBJECT:

Instructional Segments Combined into Units.

ISSUES/QUESTIONS:

1. Effectiveness of the unit
   a. To what degree do students attain mastery of the desired behaviors as a result of the instruction?
   b. Which segments or portions are ineffective in producing desired outcomes? In what ways?
   c. How could the sequence of the segments be improved?
2. Efficiency of the units
   a. How much student time (mean and range) is required to complete the unit?
   b. Which segments could be shortened or eliminated without unduly affecting the learning outcomes?
3. Student reactions to the experience
   a. What do students perceive the strengths and weaknesses of the unit to be?
   b. What proportion of the students reacted negatively to the experience? What were there reactions? What suggestions do they have for improving the unit?
   c. How complete and authoritative do students perceive the materials to be?

INSTRUMENTATION:

Mastery tests
Attitude scale

Checkpoint #5

Operational Monitoring and Evaluation

EVALUATION OBJECT:

The Operation of the Training System and its Components

ISSUES/QUESTIONS:

1. Materials
   a. What course materials are lacking?
   b. Which course materials seem to be inadequate? In what ways?
2. Equipment
   a. What necessary equipment and training devices are lacking?
   b. What equipment is in need of repair or replacement?
3. Physical facilities
   a. What limitations or inadequacies exist in the physical facilities which hamper the functioning of the system?
4. System charateristics
   a. Which components in the system interfere with or have a negative effect upon other components?
   b. What common or recurring problems are confronted in attempting to use the system in the prescribed manner?
   c. What local adaptations/modifications to the system have been made?
   d. In what ways are the plans for implementing the system inadequate? How can they be improved?
   e. In what areas does the system need greater flexibility?
5. Time schedule
   a. Are students progressing through the course according to schedule? Where do they appear to need more or less time?
   b. What local constraints limit the flexibility of your schedule?
6. Staff
   a. What staffing problems exist? What roles or assignments need to be changed?
   b. In what specific areas do the instructors need additional training?
7. Student morale
   a. To what extent do the students exhibit positive or negative attitudes towards the course or any of its components? What do they like and dislike most about it?
   b. To what extent do the students feel they are making progress towards the goals of the course? To what degree do they view the course goals as being worthwhile?

INSTRUMENTATION:

Checklists
Interview schedules

Checkpoint #6

EVALUATION OBJECT:

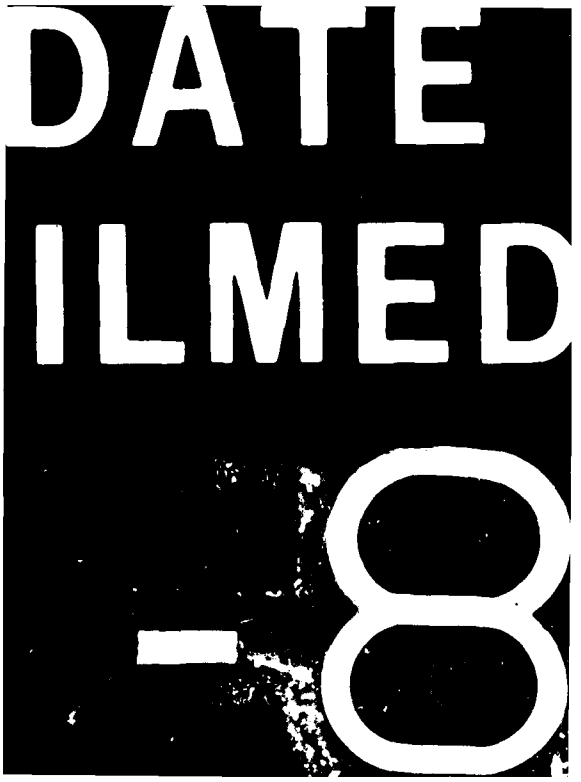
The Complete F-16 RTU Course

ISSUES/QUESTIONS:

1. Deficiencies in pilots' performance caused by inadequacies in the course
   a. What specific skills do the graduate pilots lack?
   b. What careless habits or errors of judgment are frequently manifest by the graduate pilots?
2. Reactions of the graduates to their RTU experience
   a. Which tasks do the graduates believe need more (or less) emphasis in the RTU course? Which, if any, parts of the course do they consider to have been a waste of time?
   b. To what extent do the graduates lack confidence in their ability to operate the F-16 aircraft and weapon system as a result of perceived inadequacies in the RTU course?
   c. To what extent do the graduates have a positive attitude toward their RTU experience? What aspects of the course do the students dislike? Why? What suggestions do they have for correcting these problems?
   d. From the students' viewpoint which performance standards are too strict or too lenient? In what sense?

INSTRUMENTATION:

Gradeslips
STAN/EVAL checklists and tests
Interview schedules


DATE
ILMED
-8